

AD \_\_\_\_\_  
(Leave blank)

**Award Number:**

W81XWH-11-C-0046

**TITLE:**

Deployable Pan-Flavivirus Assays for Screening Pools of Medically  
Relevant Arthropod"

**PRINCIPAL INVESTIGATOR:**

John Cooper

**CONTRACTING ORGANIZATION:**

CustomArray, Inc.

Bothell, WA 98011

**REPORT DATE:**

September 2011

**TYPE OF REPORT:**

Final Phase I

**PREPARED FOR:** U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

**DISTRIBUTION STATEMENT:**

Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</small>		
1. REPORT DATE (DD-MM-YYYY) 12-09-2011	2. REPORT TYPE Final Phase I	3. DATES COVERED (From - To) 02-14-2011 to 09-13-2011
4. TITLE AND SUBTITLE  "Deployable Pan-flavivirus and Pan-alphavirus Assays  for Screening Pools of Medically Relevant Arthropod"		5a. CONTRACT NUMBER W81XWH-11-C-0046
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)  John Cooper jcooper@customarrayinc.com		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  CustomArray, Inc.  Bothell, WA 98011		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S Army Medical Research And Materiel Command Fort Detrick, Maryland 21702-5012		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION / AVAILABILITY STATEMENT  Approved for Public Release; distribution unlimited		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT CustomArray produces a field-deployable electrochemical microarray platform that is ideally suited to the multiplexed detection of pathogenic organisms. In our phase I effort we prototyped a complete detection process which includes: Probe design, sample prep, multiplexed PCR amplification, hybridization, detection and data analysis. We used our bioinformatic system called Genotyper to create an updated pan-flavivirus and pan-alphavirus detection system. We prototyped a detection system that relies on purification of viral nucleic acids directly from arthropods, primer directed amplification of these nucleic acids, hybridization of this material to microarray bound probes, detection of hybridization patterns, and finishes with an analysis that is able to accurately type the amplified material. Our pilot studies demonstrated the detection of 5 viral species, 2 Flaviviruses and 3 alphaviruses, with near single copy detection in a single reaction/detection assay. The detection device itself is hand-held and requires only a USB connection as its sole power source. Because we can use up to 2240 probes per assay, we are able to offer a resolution and redundancy that is unprecedented in a compact, field-deployable device, and will be able to expand our assay to all medically relevant arthropod-borne viruses.		

<b>15. SUBJECT TERMS</b> Multiplexed Pathogen Detection, Hand-held MicroArray Detection, Field-deployable, Multiplexed biosensor, lab on a chip					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U	UU	34	USAMRMC
					<b>19b. TELEPHONE NUMBER</b> <i>(include area code)</i>

Table of Contents

	Page
Introduction.....	5
Body.....	6
Key Research Accomplishments.....	31
Reportable Outcomes.....	32
Conclusion.....	33
References.....	34
Appendices.....	34

## **Introduction**

There is a pressing need to detect pathogens, invasive species, and biothreat agents using a compact and robust assay device. By bringing robust and multiplexed detection into the field, we will be able to dramatically improve the discovery of emerging pathogenic threats. The CustomArray system is the only system that has the breadth and the redundancy to accurately detect multiple species in real world situations where novel forms are continually emerging. The purpose of this phase I SBIR research effort was to develop a prototype assay of a field-deployable electrochemical microarray assay for the multiplexed detection of arthropod-borne alphaviruses and flaviviruses. The scope of this effort encompassed the research necessary to build a complete detection process, which included: Probe design, sample prep, multiplexed PCR amplification, hybridization, detection and data analysis. The tasks performed during the phase I period, which were necessary to create and test this prototype assay, are outlined in the next section.

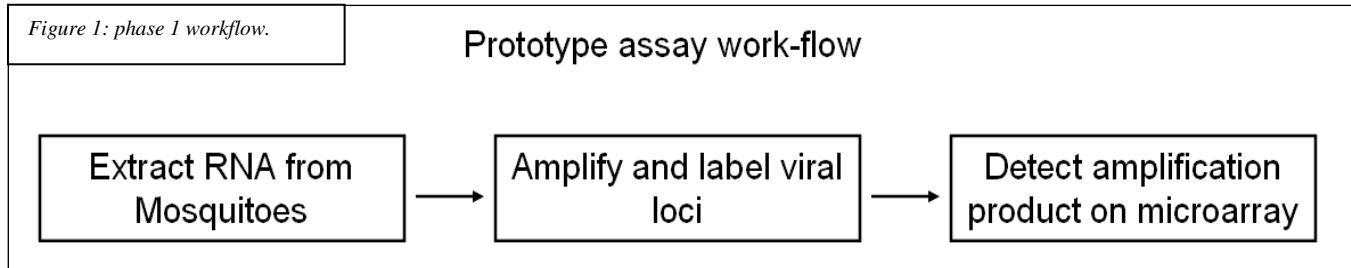
## BODY

Our Phase I research effort resulted in a prototyped detection system that relies on the following steps: (1) Purification of viral nucleic acids directly from arthropods, (2) Primer directed amplification of these nucleic acids, (3) Hybridization of this material to microarray bound probes, (4) Detection of hybridization patterns, and finishes with (5) Analysis that is able to accurately type the amplified material. Our pilot studies demonstrated the detection of 5 viral species, 2 flaviviruses and 3 alphaviruses, with near single copy detection in a single reaction/detection assay. The tasks to develop this prototype assay are listed below:

### Tasks:

1. **Bioinformatics:** Viral genomic loci were chosen for amplification. From these loci, we designed multiplexed PCR primers for amplification of these loci and probes for interrogating the amplified material.
2. **Primer Testing & Alternative amplification testing:** Amplification primers were tested and validated on pure genetic material from individual viral genomic samples, which included 3 alphaviruses and 2 flaviviruses. The individual PCR reactions for each of these 5 test viruses were multiplexed into one PCR reactions. The sensitivity and multiplexability of this amplification scheme was then characterized. In parallel to the multiplexed PCR detection, we explored an alternative amplification method: a padlock assay.
3. **Chip design testing:** Amplified labeled DNA amplicons were hybridized to the first version of the chip, which was designed from the results of our bioinformatics task. Probes from this chip design were chosen based on the specificity and signal strength of these hybridizations. These selections were used to generate a second and third version of the chip design which resulted in a more compact design using a few hundred probes per virus tested. We expect that 2240 probes will be enough to cover all the sequence space needed to type all the viruses listed in the FOA. Although we only tested 5 organisms during the phase I period, the chip design synthesized and used for testing included all of the viruses in the FOA.
4. **Sample Prep:** We evaluated methods for grinding and extracting genetic material from crushed arthropods.
5. **Data Analysis:** A prototype data analysis pipeline was written to analyze multiple hybes in batch mode. This software can analyze the hybridization patterns and make a call based on probe signal intensities.
6. **Deliverable:** Provided 100 assays, along with a detection device and protocol to the COR.

The results of these tasks created a prototype assay for arthropod-borne viral detection which is succinctly summarized in the work-flow outlined in figure 1 below:



The completion of these tasks was integral in developing this prototype detection assay and device. As such, a summary of the work completed for each task is reviewed below:

## Task 1: Bioinformatics: Probe and Primer Design

### Overview:

The overall goal of this project was to design an assay for the selection of viruses from class A arboviruses (alphaviruses) and class B arboviruses (flaviviruses). In our bioinformatics effort, we designed the primers and probes that were used to amplify and hybridize the biological sample. To this end, it was necessary to recode the probe design software system to handle the large amount of sequences that were analyzed during the design phase. We used this probe design system to create the primers for PCR amplification as well as the detection probes. A summary of the bioinformatics tasks for the probe and primer design are outlined below:

1. Chose Locus for Probe design for both Alphavirus nsP4.
2. Chose Locus for Probe design for both Flavivirus ns5.
3. Built Genus-level sequence database for Alphavirus with 4554 accession numbers
4. Built Genus-level sequence database for Flavivirus with 32242 accession numbers.
5. Chose accessions and gene loci to be used as input for probe design system.
6. Adapted Probe design system to design and choose probes for probe design.

### Completed Bioinformatics Tasks: Create Probe/Primer Design Database

In the first phase of this project, we designed probes that are specific for these organisms at the genus level. We first chose loci that are easily amplified and well characterized in these two classes of organisms. These loci were used as input for the design process. The design process consists of several modules and compiled libraries of code. These modules implement those algorithms used to design, test and choose well-behaved hybridization probes. These probes were characterized based on their specificity against a database of related sequences. These related sequences were organized based on the genus-species and consist of all the available sequences of the alpha and flavi-virus families. Probe specificity was, thus, designed to reflect not just the small subset of viruses that the user is interested in, but also the entire list of available sequence. In this way, these probes were designed to be compatible with any conceivable addition or subtraction of detection probes.

While the input sequences for probe design can be limited to a few well characterized sequences, the database against which we compared these probes was chosen to contain as much of the available information possible. All the available sequences in genbank for Alphaviruses and flaviviruses were downloaded and organized into clusters based on the species of the source organism. Probes were blasted against this database, and their specificities were reported against this database. Each probe's specificity was reported against a vector each of whose elements correspond to all the relevant species within these viral families. In this way cross-reactivity, if it existed, could be evaluated across the entire sequence space that was sampled by the assay. Figure 2 summarizes the compilation of alpha and flavi-virus families for probe and primer design.

**Figure 2: Compilation of alpha and flavi-virus families for probe and primer design.**

alphaviruses		taxID	rank
Chikungunya	SFV Complex	37124	species
O'nyong-nyong	SFV Complex	11027	species
Sindbis	WEEV complex	11034	species
Venezuelan equine encephalitis virus	VEEV complex	11034	species
Mayaro	Mayaro virus	59301	species
flaviviruses			
	Dengue	12637	species
	Dengue1	11053	norank
	Dengue2	11060	norank
	Dengue3	11069	norank
	Dengue4	11070	norank
tick-borne encephalitis virus complex		29263	subgenus
	Tick-borne encephalitis virus	11084	species
	Kyasanur Forest disease	33743	species
	Russian spring-summer encephalitis	64316	species
	Far Eastern tick-borne encephalitis		
	Central European tick-borne encephalitis		
	Omsk hemorrhagic fever virus	12542	species
	Japanese encephalitis virus	11072	species
	West Nile virus	11082	species

## Completed Bioinformatics Tasks: Probe Design

The first version of our chip design included 1861 probes which were created by targeting 15 species of arboviruses from both flavivirus and alphavirus families. These probes are summarized in figure 3.

To complete this probe selection, all available alpha and flavivirus sequences were collected from Genbank using the NCBI tax browser. These sequences were used to determine probe specificity. Each sequence was placed in a category based on its genus. These were combined to form a list of all available alphavirus sequences, labeled by their species. For each species we were interested in, sequences (15 in all) were collected, genbank records were parsed and sequence regions containing the genes of interest (ns5 for Flavivirus and ns4 for alphavirus) were collected. For each family of

**Figure 3: summary of probes for chip design version #1**

Species	Probes Designed	
AV-CHIKV	140	Semliki Forest Virus Sindbis
AV-MayaroV	44	
AV-OnyongnyongV	155	
AV-SFV	55	
AV-SV	198	
AV-VEEV	198	
FLV-DengueV1	188	
FLV-DengueV2	100	
FLV-DengueV3	130	
FLV-DengueV4	121	
FLV-Japanese encephalitisV	200	
FLV-Kyasanur forest diseaseV	75	
FLV-Omsk hemorrhagic feverV	120	
FLV-Russian Spring Summer encephalitisV	37	
FLV-West NileV	100	
	1861	



sequences, these regions were aligned and extraneous regions were trimmed in using the program Jalview. The most conserved alignment block was extracted and used for probe design. A consensus sequence was generated from each alignment, and often this sequence could be used in other species to find the orthologous gene regions.

After each design set was generated, probes were designed by tiling each sequence. Probes were designed to be between 25-35bp, with a melting temperature ( $T_m$ ) of 72C. Probes exhibiting more than 6 repeat bases or two-base repeat structures were rejected. So were probes with hairpin structures with computed melting temperatures > 65C. Finally, probes were designed to have a GC content between 35-65%.

After probe design and selection, each probe was blasted against the large arbovirus database generated previously. In order to select probes, probes are sorted in several steps. In the first place they are selected by increasing numbers of other species Hit, then inversely by the number of hits. This ranks probes based on specificity first and then by penetrance second. These sorting routines are automated within our probe design software; however, we checked the probe selection process manually. Furthermore, for a few instances, we made sure that the regions hit corresponded to unique regions within the alignment.

## Completed Bioinformatics Tasks: Primer Design

For the development of the phase I prototype, we narrowed the assay to include detection of three alphaviruses and two flaviviruses, as outlined in the FOA. The alphaviruses included Chikungunya virus, Sindbis Virus and Venezuelan equine encephalitis virus (TC-83 strain). The flaviviruses included West Nile Virus and Dengue 3 virus. We chose these organisms because these five species were made available to us by our COR: Dr. John Lee, Virology Division, USAMRIID. Please note that all species are considered non-select agent viruses/nucleic acids in Qiagen AVL buffer.

For designing primers for amplification of the alphaviruses, we focused on the Non-structural protein 4(nsp4) which is a member of the RdRP\_2 [RNA dependent RNA Polymerase] superfamily, it is conserved for most alpha viruses. For the Flaviviruses, we focused on the non-structural 5 protein (which is the functional equivalent of alphavirus nsp4). These gene loci were chosen because these specific proteins have been the primary focus of phylogenetic study in these viral families. For this reason, there are many isolates for this gene family available in Genbank. To choose primers specific for these regions, we used a primer design routine that we have designed previously. This system takes a block of sequences and designs all possible valid PCR primers that conform to the same quality metrics as were applied to the probes, with a few changes: Probes were designed to be between 18-25bp, with a melting temperature ( $T_m$ ) of 60C. Probes exhibiting more than 6 repeat bases or two-base repeat structures were rejected. So were probes with hairpin structures or Palindromes with computed melting temperatures > 65C. Finally, probes were designed to have a GC content between 35-65%. Forward primers were written in sense orientation. Reverse primers were written in the anti-sense, and bases were added to the 5' end such that the  $T_m$  was adjusted up to 72C. In this way, the melting temperature of the primers after a few cycles of amplification became imbalanced by about 10C. This is done so that single stranded DNA synthesis can be performed in the later cycles. At these cycles, the annealing temperature is raised from 60C to 72 C for additional 15-20 cycles. This generates reverse single stranded DNA, that is suitable for hybridization to microarray-bound probes.

## Summary of Bioinformatics effort; significance for our Phase II effort:

The bioinformatics effort completed in this phase I period lay the ground work for expanding the assay to include all medically relevant arthropod-borne viruses. Using the database that we created along with the optimized design algorithms, we will be able to expand our prototype assay “in silico” to include all of the viruses listed in the FOA as well as others that may prove relevant, ultimately depending upon the field setting. This bioinformatics effort will, in the end, be the cornerstone for expanding the assay in the phase II period.

## Task 2: Primer Testing & Alternative amplification testing

### Primer testing:

Our strategy for the amplification of the viral RNA material for detection on a microarray is as follows:

1. Reverse transcription with specific primers
2. 40 rounds of PCR (3 temperature cycling)
3. finish with 30 rounds of a high temperature PCR (2 temperature cycling)

We used SuperScript III One-Step RT-PCR with Platinum TAQ (#12574-026) for the RT-PCR amplification, and we used biotinylated dCTP to integrate a biotin label into the amplified material. The RT-PCR reaction was performed in 25ul reaction volumes by combining the following reagents:

1. dH2O	6.3 ul
2. 2x RXN buffer (Invitrogen)	12.5ul
3. 5mM MgSO4 (Invitrogen)	2.0ul
4. 0.4 mM Biotin-14 dCTP (Invitrogen)	0.7ul
5. Primer Pools (10uM each)	2.0ul
6. Enzymes (Invitrogen)	0.5ul
7. ++sample (purified RNA)	1ul

The reactions were mixed into PCR tubes and incubated on a thermal cycler using the following temperature cycling scheme:

Step 1:	50C	30 min
Step 2:	94C	4 min
Step 3:	94C	30 sec
Step 4:	55C	45 sec
Step 5:	72C	45 sec
Repeat step 3-5, 39times		
Step 6:	94C	30 sec
Step 7:	68C	60 sec
Repeat step 6-7, 29times		
Step 8:	94C	5 min
Step 9:	4C	hold

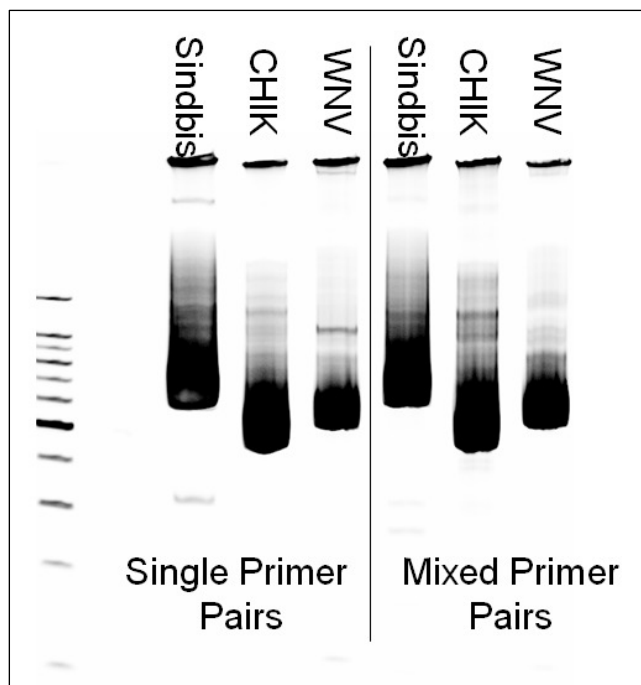
We received viral lysates for the following species: Alphaviruses: Chikungunya (CHIK), Sindbis, and Venezuelan Equine Encephalitis (VEE, TC-83 strain). Flaviviruses: Dengue 3 and West Nile Virus (WNV). We purified these lysates using the QIAamp Viral RNA Mini Kit as per the manufacturer’s protocol. We were provided titer

information for three of the five viruses, which were Sindbis, CHIK, and WNV. We were not provided with titer information for Dengue 3 and VEE (TC-83). Assuming that we purified the lysates with 100% efficiency (extremely conservative estimate), and given the titer information provided, we can calculate the concentration of viral genomes for Sindbis, CHIK, and WNV to be  $1.2 \times 10^5$ ,  $5.6 \times 10^5$ , and  $4.9 \times 10^5$  pfu/ul, respectively.

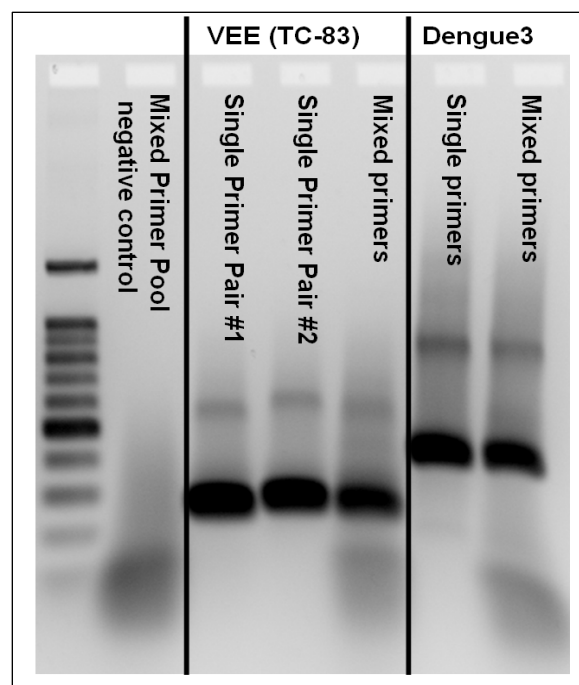
To review: primers were designed to be between 18-25bp, with a melting temperature ( $T_m$ ) of 60C. Forward primers were written in sense orientation. Reverse primers were written in the anti-sense, and bases were added to the 5' end such that the  $T_m$  was adjusted up to 72C. In this way, the melting temperatures of the primers after a few cycles of amplification become imbalanced by about 10C. This is done so that single stranded DNA synthesis can be performed in the later cycles. At these cycles, the annealing temperature is raised from 55C to 72 C for additional 15-20 cycles. This generates reverse single stranded DNA, which is suitable for hybridization to microarray-bound probes.

Primers were tested separately as primer pairs specific for each individual virus, and mixed together as a primer mix of all primers for each of the five viruses. The primer pairs and mix were tested on each virus, individually. Figure 4 shows results of this testing from the amplification reaction for Sindbis, CHIK, and WNV. Aliquots from the RT-PCR amplification reaction were removed at cycle 24 of the first PCR cycle run out on a 6% polyacrylamide gel. In a separate experiment, Dengue 3 and VEE (TC-83) were tested in a similar fashion. VEE had the added complication of needing two primer pairs to amplify the chosen loci. At the end of the RT-PCR amplification reaction, aliquots were removed and run out on a 2% Agarose Gel. These gels, figures 4 and 5,

**Figure 1: Sindbis, CHIK, and WNV primer validation**



**Figure 5: Dengue 3 and VEE primer validation**

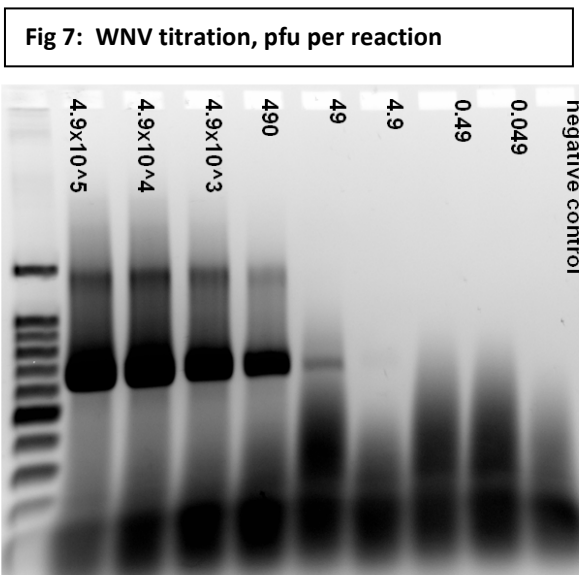
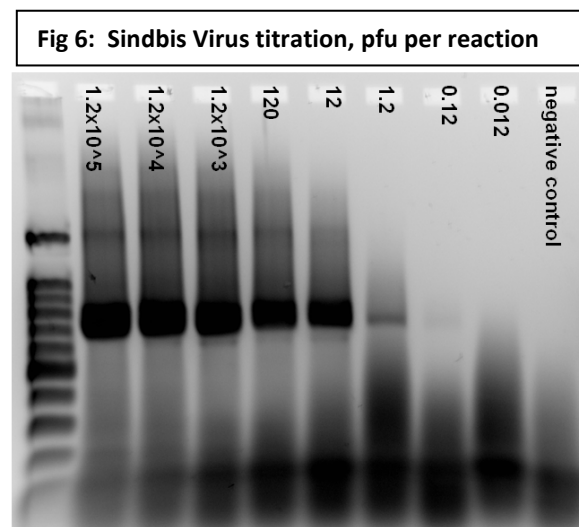
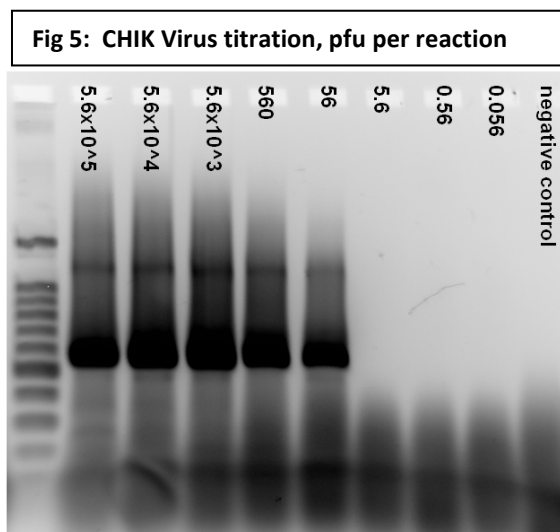


summarize the primer validation: All 5 species were successfully amplified, individually, using the primer mix specific for all 5 species: Dengue3, VEE, WNV, Sindbis, and CHIK. With these primers, we did not observe problems with primer-dimers or inhibition of the specific PCR reaction. The next test was to take this amplified

material and hybridize it to our version 1 microarray chip. These data are discussed in “Chip Design Testing” task section.

## RT-PCR amplification testing and validation:

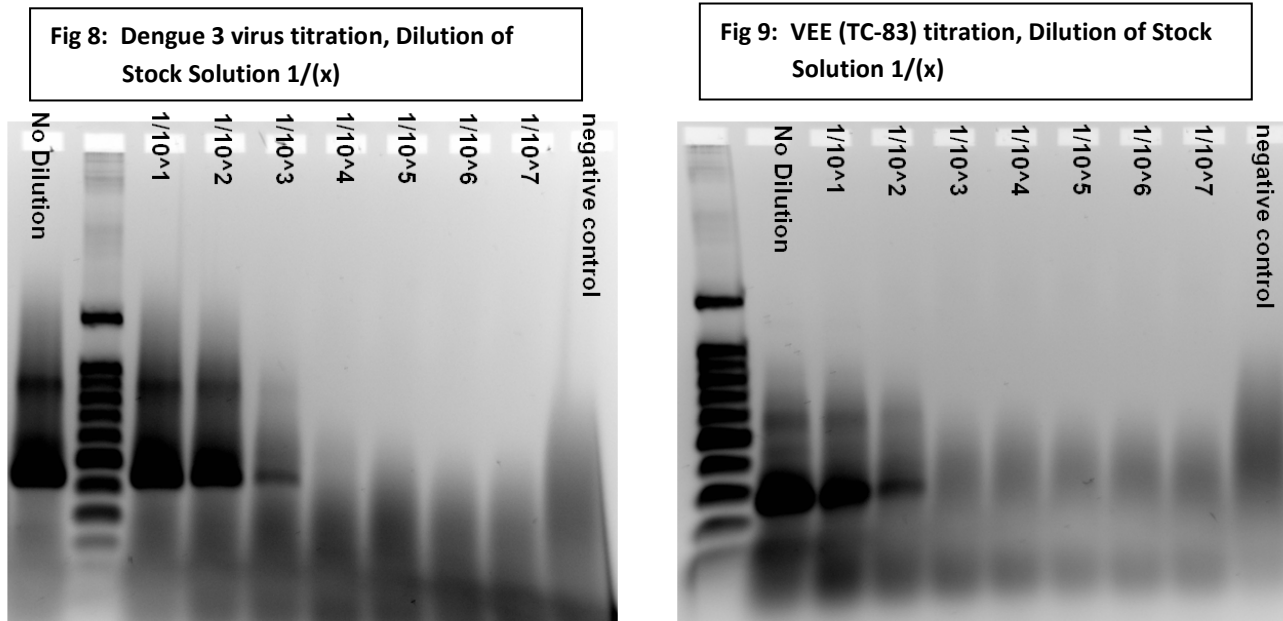
To explore the lower limits of detection of our RT-PCR/ssDNA amplification strategy, LOD titrations were performed on each specie sample. We started our LOD titrations with Sindbis, CHIK, and WNV. The titer of these viral lysates were provided and we estimated (assuming 100% purification efficiency) that the concentration of viral genomes is  $1.2 \times 10^5$ ,  $5.6 \times 10^5$ , and  $4.9 \times 10^5$  pfu/ul, respectively, in our viral RNA stock solution. In a serial extinction dilution series, we tested 8 concentrations of each virus in a separate RT-PCR amplification reaction. We started with 1ul of the stock solution and made 7,  $1/10^{\text{th}}$  serial dilutions. In this way we would be sure to have a dilution which would not contain any viral genome. Results of these titrations are shown in figures 5-7:



In all of these titrations, we are clearly approaching single copy detection. Furthermore, The Sindbis and WNV LOD titrations were hybridized to version #1 and version #3 microarray designs. Those results, which are shown in the next section, show hints of being able to detect the virus when it is no longer presenting a visible band on the gel. In the case of WNV, we appear to be able to detect 4.9 copies of the viral genome on the array, which is clearly NOT visible on the GEL image to the right.

We continued testing the LOD of both VEE (TC-83), and Dengue 3. We do not have information as to the titer of

these samples. Nonetheless, we preformed an identical dilution series as described above from our RNA stock solutions. These results are displayed in figures 8 and 9 below:



The Dengue titration above was hybridized to a version #2 microarray, and showed clear detection of Dengue3 at 1/10<sup>4</sup> dilution condition, which is NOT visible above. Furthermore, the 1/10<sup>5</sup> condition was NOT detected on the array (refer to reporting period #5 for a detail of these array results).

Thus far, we have demonstrated that our RT-PCR strategy for producing hybridizable material in a multiplexed fashion is both specific and sensitive. Using this strategy we demonstrated near single copy detection in LOD tests of each of the three viruses with which we have titer information: Sindbis, CHIK, and WNV. However, all of these amplification reactions were detecting the presence of a single virus, that is, only one viral type was present for amplification in the RT-PCR reaction. The ability to detect multiple pathogens, if present, would be a necessary component of any field deployable test. For this reason we further defined our amplification strategy to verify the co-amplification of multiple viral species.

To test the RT-PCR reaction for co-amplification, we carried out our normal amplification strategy of RT-PCR for 40 cycles, followed by 30 rounds of a high temperature PCR (2 temperature cycling) to produce single stranded product for hybridization to the microarray. Four conditions were tested:

- Condition #1: Sindbis, CHIK, and VEE (All Alphaviruses)
- Condition #2: WNV and Dengue3 (All Flaviviruses)
- Condition #3: WNV and VEE
- Condition #4: Sindbis, CHIK, and Dengue3

For the initial test, 1ul of viral stock solution was used for each virus. The RNA stock solutions for Sindbis, CHIK, and WNV were estimated, using titer information provided, to be  $1.2 \times 10^5$ ,  $5.6 \times 10^5$ , and  $4.9 \times 10^5$  pfu/ul, respectively. Using the RT-PCR LOD titrations of VEE and Dengue3 shown in figures 8 and 9, we can estimate a “rough” pfu concentration of these stocks to be  $\sim 10^3$  and  $\sim 10^4$  pfu/ul, respectively. The RNA stocks were combined in one 25ul RT-PCR reaction as per the conditions outlined above. 7.5ul of the final reaction was run out on a 2% Agarose gel, shown on the left in figure 10 below as “co-amplification#1”. Conditions #2 and #3 showed successful co-amplification reactions in that the appropriate viruses were amplified and display the correct sized PCR fragments. Notably, conditions #2 and #3 were co-amplifying only 2 viral species. Conditions #1 and #4 were only partially successful, and, notably, were tests to co-amplify 3 viral species. Condition #1 amplified Sindbis and CHIK, but failed to amplify VEE. And condition #4 amplified Sindbis and CHIK, but failed to amplify Dengue3. The reason for the failure to amplify all products in these conditions isn’t entirely clear from this experiment. Since the RNA stock concentrations of both VEE and Dengue3 are the lowest, one possibility is that the viruses at higher concentration are amplified first, resulting in a plateau of the RT-PCR reaction and an inability to amplify those species at lower concentrations. To test this hypothesis, we performed another experiment, “co-amplification #2”, where the same conditions, conditions 1-4 above, were replicated using adjusted RNA stock input quantities for the RT-PCR reaction. Sindbis, CHIK, and WNV were diluted to 1/50<sup>th</sup> stock concentration while VEE and Dengue3 were undiluted. 1 ul of these solutions were then used as input into the RT-PCR reaction, and those results are shown at right in figure 10.

The results show all conditions amplifying all viral species that were present. Furthermore, this gel image adds strength to our hypothesis that the RT-PCR reaction may be consumed by those viral species which are at higher concentrations in the RNA sample that is tested. Whether this will be an issue in actual field testing is unknown and will have to be determined by empirical testing of realistic samples.

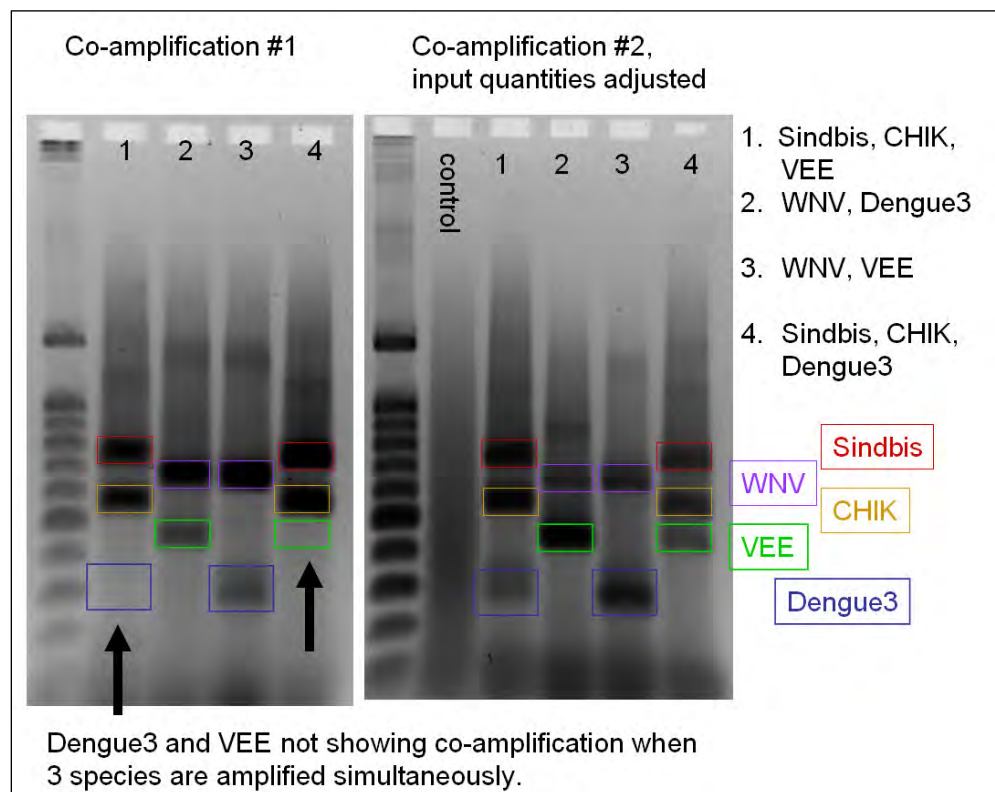


Figure 10: RT-PCR validation in co-amplification reaction conditions.

## Alternate

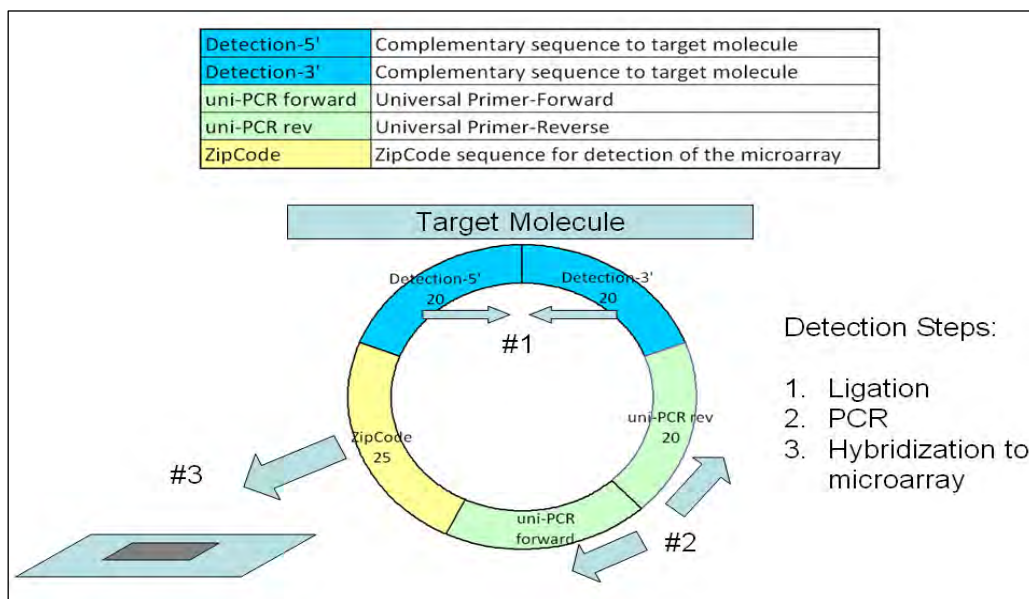
## amplification testing, Padlock Detection:

During the phase I period, we explored alternative amplification strategies, other than RT-PCR, for our prototype assay. We anticipated that there could be problems with our multiplexed RT-PCR strategy from increased primer-primer interactions when the PCR multiplex is increased. Ultimately, we saw no problems in our RT-PCR reaction when multiplexed to detect 5 viruses, but problems may occur when multiplexing to detect a higher number of viruses. Of course, our planned strategy to circumvent this issue is to simply have multiple low complexity multiplexed RT-PCR reactions. None the less, we explored alternate amplification techniques to devise a methodology to have one amplification reaction.

During this phase I period, we piloted experiments to test the feasibility of using a variation of the MIP assay to detect an RNA target, called Padlock Probe detection. The multiplexed MIP, or molecular inversion probe, assay is a potential method for replacing multiplexed PCR by a universal PCR step. Briefly, MIP probes are linear DNA probes with 5' and 3' terminal ends that are complementary to adjacent sequences on the DNA template to be detected. When the two ends of the probe hybridize to their complementary target sequences, the probe circularizes leaving a gap for a DNA polymerase to fill-in, which effectively copies the template, much like a PCR. After the extension reaction, a Ligase ligates the ends together, circularizing the inversion probe. In the middle of the MIP sequence is a universal primer pair for subsequent amplification. Only when the MIP is successfully circularized will the universal PCR successfully amplify the copied template. Because of this specificity, MIP probes can be highly multiplexed. Each MIP probe can be considered a micro-reactor for PCR because the flanking recognition sequences are physically tethered together. Furthermore, a universal PCR can effectively amplify multiple MIPs simultaneously.

A simpler version of this assay is to use the MIP as a padlock probe, where circularization requires only a ligation to occur. This has advantages in a more streamlined assay; however, each probe will need to be specific for each virus. This assay is described in the figure 11 below:

Figure 11: Padlock probe design and assay work-flow



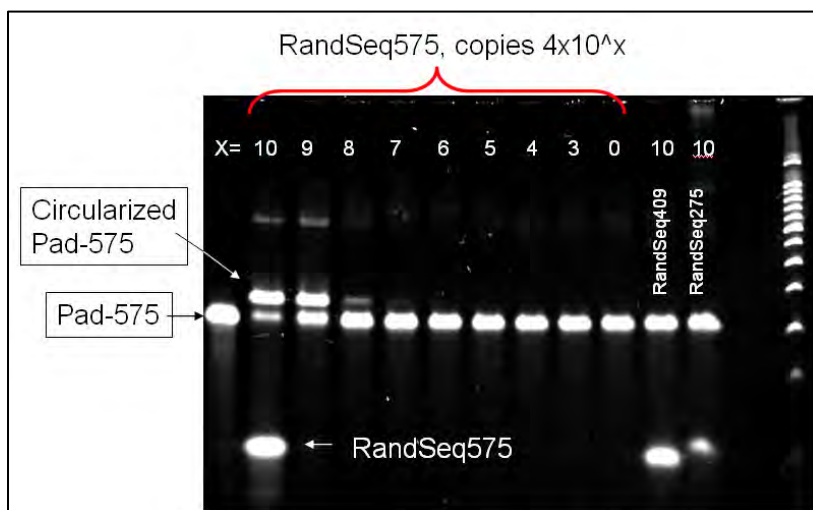


Briefly, the assay seeks to use a padlock probe whose 3' and 5' ends hybridize to a target sequence bringing the ends of the probe together. Using a template-dependent ligase, such as Taq DNA Ligase, the 3' and 5' ends of the probe are ligated and thus circularized. Once circularized, the probe serves as the template for a PCR reaction. The PCR reaction will amplify a zip-code sequence present in the circular probe as well as the target sequence ONLY if a ligation occurs. Both of these sequences can be used to detect positive amplification via hybridization to our array and subsequent electrochemical detection.

Using the Padlock detection scheme described above, we piloted an experiment to detect a test RNA sequence. The RNA sequence was generated randomly, and we designate it as RandSeq-575. We then designed a

complementary padlock probe, named pad-575, containing complementary sequences to this RandSeq-575. To test for successful detection of RandSeq 575, we piloted an experiment where decreasing concentrations of the RNA target were incubated with a static concentration of Padlock probe. The ligation, using TAQ DNA Ligase, was carried out as per manufacturers protocol (NEB) at 45C for 30 minutes. The figure below shows a ligation product gel featuring the pad-575 + RandSeq-575 model. The gel at right, figure 12, shows circularized padlock, non-circularized padlock, and the target (when visible due to concentration). Dilutions were 10 fold starting at  $4.0 \times 10^{10}$  and ending in 4000 copies. The last three lanes (excluding 100 BP ladder) contain a zero target, and high point concentrations of 2 different incorrect targets, randSeq409, and randSeq275.

Figure 12: Padlock probe pilot experiment: detection of synthetic RNA sequence



The condition containing  $4 \times 10^6$  copies and 0 copies of RandSeq575 were subsequently used as template for a universal PCR reaction as positive and negative controls, respectively. The PCR reactions were carried out with 30 cycles of thermal cycling using Taq polymerase. The forward universal primer used in these PCR reactions was biotinylated. In this way, the subsequent PCR products would have a biotin label, enabling Electrochemical Detection on our microarray. These RAW PCR products were each mixed 1:1 with 6xSSC. 30ul of each of these PCR mixes were incubated for 30 minutes in separate sectors of a 4x2k microarray with detection probes complementary to the forward priming product. The chips were washed and incubated with Streptavidin-HRP for 15 minutes, washed and run through the standard CustomArray electrochemical detection protocol. The hybridization results (not shown, refer to reporting period 2 results) showed specific detection of the amplification product in the microarray assay.

Continuing from this proof of principle amplification scheme, we designed Padlock probes specific for each of the 5 viruses. The design of these probes was based on our microarray hybridization results from the version #2 and version#3 chip design hybridization data. This hybridization data shows which probes hybridize with specificity to each virus. Our design of the probes was as follows: For each virus, the highest binding probes



(probes with the highest intensity) were selected as padlock candidates. These probes were cross-referenced to the hybridization results from the other viruses. If the other viral hybes showed cross-hybridization to the candidate padlock probes, then they were excluded. We set a cut-off threshold of the mean value of each hybridization (the average value of all probes in the hybridization), and excluded candidate probe which had values above this threshold in any one of the other viral hybes. In this way we believe that we have selected specific padlock probe candidates. Unfortunately, we were unable to complete a fully functional assay with the designed padlock probes during this phase I effort. In our Phase II assay work, we will design an instrument that will be able to accept multiple different amplification techniques, which will lower the overall risks to the final assay, and make the final instrument flexible for applications where RT-PCR is not appropriate. In this way we hope to co-validate this assay along with the RT-PCR assay in our phase II work.

## Task 5: Data Analysis

Before discussing the validation of the microarray hybridization detection assay, we describe here our work to create a data analysis system to analyze the hybridization results. As part of the final deliverable for this Phase I SBIR, we developed a data analysis pipeline to analyze multiple hybes in batch mode. This software can analyze the hybridization patterns and make a call based on probe signal intensity. To this end, we have installed a “plug-in” to the ElectraSense software package as a prototype to make these calls.

Below, figure 13, is a screen shot of the ElectraSense Software, which displays three main functions: “Experiment”, “Data”, and “Results”. The “Results” function is selected in this picture. After the normal scan is complete, the user simply has to go to this function of the software and follow the directions to generate graphical data based upon the intensities of the probes for each species. Each sector of the 4x2k array will be summarized in both graphical (see tabs below: Results Graph 1-4) and numeric form, allowing for exportation into EXCEL, for example.

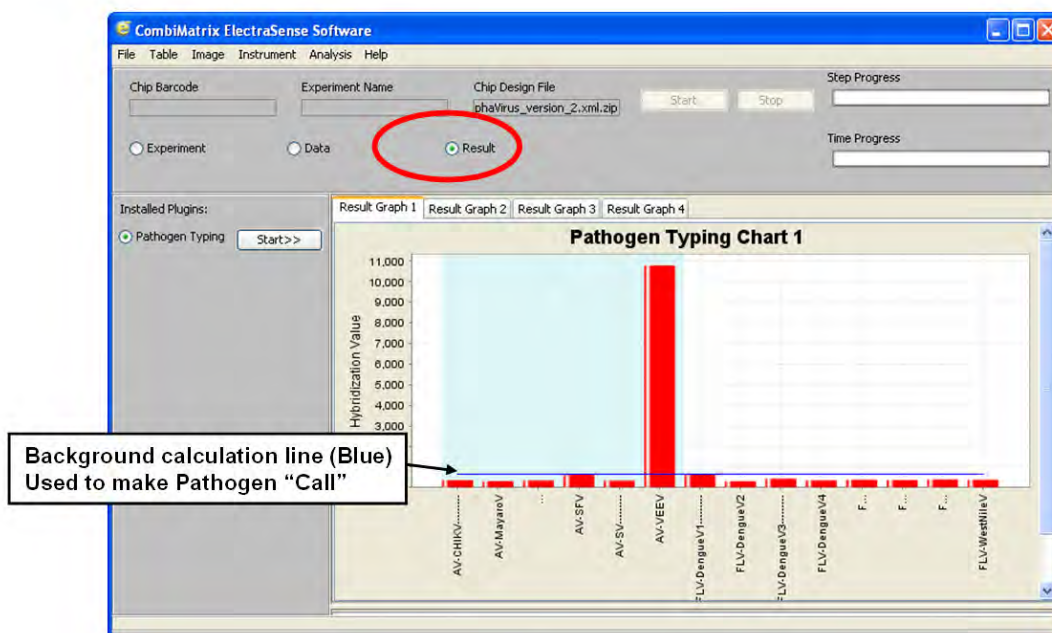


Figure 13: ElectraSense data analysis plug-in interface

The data analysis built into the ElectraSense software works as follows. On the array, there are probes associated with each virus, and the post-hybridization intensity values of these form a set of probe intensities associated with each virus. The software allows the user to select a high and low cutoff (define these as H and L, respectively) such that, for each set, the highest H percent of the values and the lowest L percent of the values are discarded. This allows the user a simple mechanism of throwing out outliers if desired. The software then computes the mean and standard deviation of each resulting set. Thus, for each virus, there is an associated mean and standard deviation value. From this, the software creates a bar graph showing the mean for each virus, and the software draws a line across the graph called the "background." The background is equal to the mean plus three standard deviations, using the mean and standard deviation of the virus that has, from among all viruses, the lowest mean. In this way, the "background" provides a means for calling the presence or absence of viral pathogens.

In the data that follows, we will use this graphical representation as well as a visual image of the actual hybridization data to show results from our chip design version #1 through #3 testing.

### Task 3: Chip Design Testing

This phase I effort resulted in the testing of 3 chip design versions, called version #1 through #3. Between each version, we refined the probes used to make the viral pathogen calls. A description of this refinement will follow. Data and visual hybridization results are shown from each version to illustrate the strength of our system in refining the probes and clarifying the resulting chip designs, which dramatically increased the specificity and sensitivity of the assay between versions #1 and #3.

Using the probes designed from our bioinformatics effort, in which 1861 probes were created targeting 15 species of arbovirus from both flavivirus and alphavirus, we populated a microarray design file to create a chip design version #1 array. The design is that of a 4x2k microarray format, so that per each microarray synthesized, there are 4 separate areas (defined by a gasketed hybe chamber) to perform 4 separate assays. Furthermore, probes from each viral specie are grouped on the array for an easy visual read-out.

To test the probe design of these version 1 microarrays, the amplified material from the PCR primer testing of Sindbis, CHIK, and WNV were hybridized to these arrays. Our detection assay is an electrochemical read-out of hybridized material. Briefly, the amplified and biotinylated material is hybridized to the array. The array is then incubated with streptavidin-HRP which binds to the biotinylated targets. The HRP facilitates the oxidation of the substrate TMB (3,3',5,5',tetramethylbenzidine) which is then detected electrochemically at the surface of each electrode.

First, we tested the Sindbis and CHIK amplification products from both the single primer pair and mixed primer pairs amplification reactions. These visual results are shown at the top in figure 14.

Second, we tested the WNV amplification products from both the single primer pair and mixed primer pair amplification reactions, as well as a mixture of all mixed primer pair amplification reactions (Sindbis, CHIK, and

WNV). These visual results are shown at the bottom of figure 14, at right.

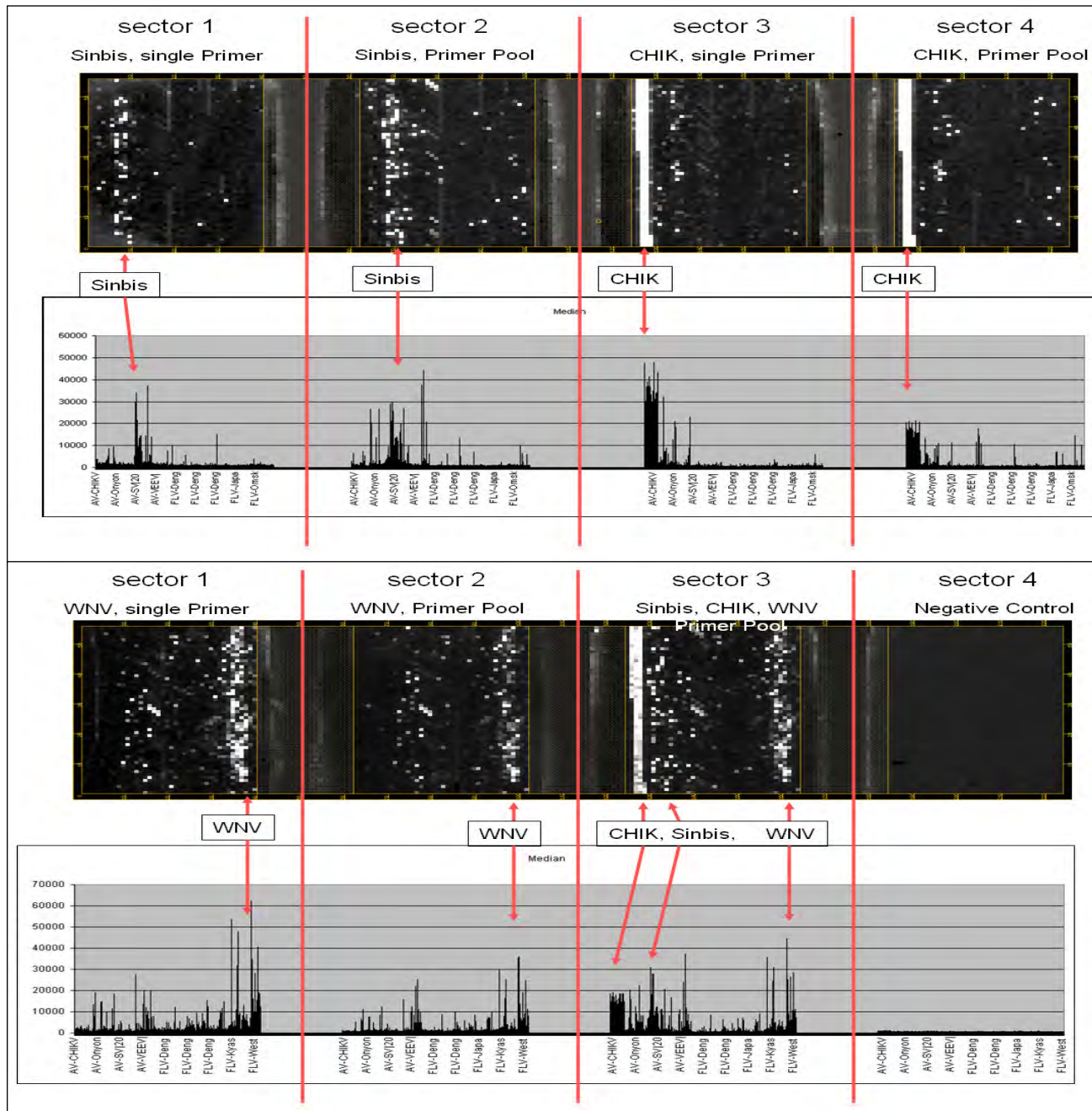


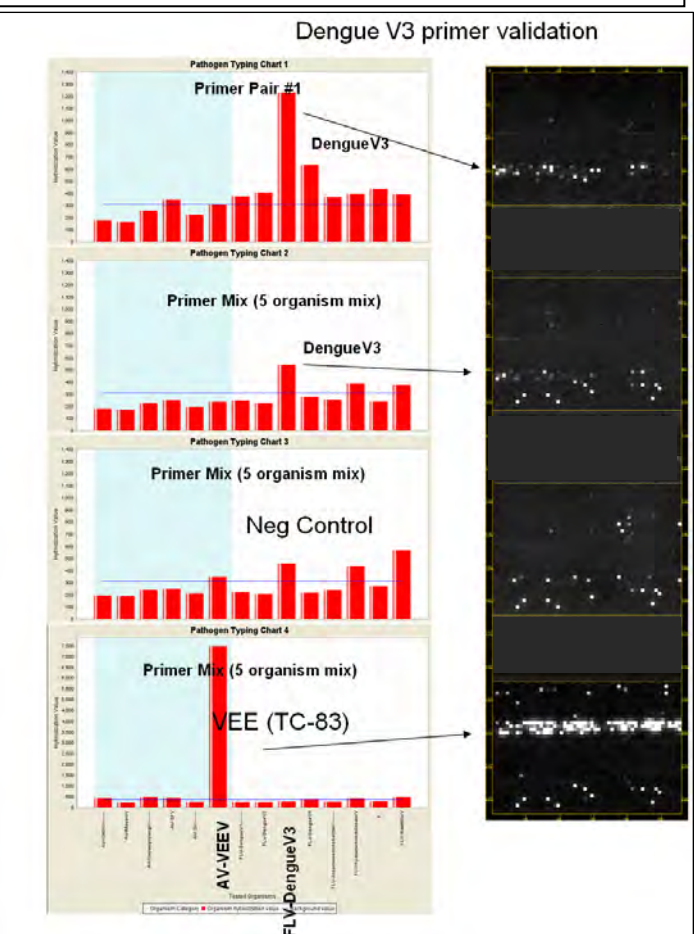
Fig 14: version #1 chip design testing on WNV, Sindbis, and CHIK RT-PCR amplification products

We continued to test the version #1 microarray chip design. The primer validation of Dengue 3 and VEE (TC-83) were hybridized to a microarray, respectively. The results from the primer validation are displayed below. To the left of each microarray images, we display the graphical analysis generated by the ElectraSense Data analysis “plug-in” (described earlier).



The hybridization of the Dengue 3 primer validation is displayed below in figure 16. Sector #1 tests the single primer pair specific for Dengue3. Sector #2 tests the total primer mix (5 organism mix). Sector #3 is a negative control, and sector #4 is a repeat of the total primer mix testing for VEE (TC-83). The hybridization results show that the probe set for Dengue is rather weak for the amplification product produced in the RT-PCR reaction with this primer pair. Nonetheless, we used these data to refine a probe set for second iteration of the microarray, version #2 microarray chip design .

Figure 16: Dengue amplification product hybed to version #1 chip



## Version #2 and #3 Chip Design Testing:

To increase the reliability of our assay, we set about refining the microarray chip design with second and third revisions. In version#2, we used the hybridization results from the version #1 microarray to select those probes which behaved well, and excluded those probes which behaved poorly, or showed high-cross hybridization. We did not design new probes for this revision. The selection of probes retained in the second version were chosen as follows: For each virus, the highest binding probes were selected from the mixed primer RT-PCR hybridization. The cut-off for this selection was set to include only those probes which were more than 1.5 times the average signal in that hybridization. These probes were then cross-referenced to the hybridization results from the other viruses. If the other viral hybes showed cross-hybridization to any of these probes, then they were excluded. We set a cut-off threshold equal to the mean value of each hybridization, and excluded probes which had values above this threshold in any one of the other viral hybes. In this way we chose probes which bound strongly and specifically to each of the 5 viruses that we have tested.

The version #2 chip design was tested against all 5 viruses in 4x2k array hybridizations. However, we omit these data here and refer to the monthly reports for a detail of those experiments.

We further refined our microarray chip design and created our final version, version #3 chip design. For this revision, we again (as with the creation of version #2) used the hybridization results from the previous microarray version, in this case version #2, to select those probes which behaved well, and excluded those probes which behaved poorly, or showed high-cross hybridization. The version #3 chip design was tested against all 5 viruses in two 4x2k array hybridizations. The RT-PCR amplification product tested on these arrays was produced from the most concentrated reactions: Sindbis at  $1.2 \times 10^5$  pfu/rxn, CHIK at  $5.6 \times 10^5$  pfu/rxn, WNV at  $4.9 \times 10^5$  pfu/rxn, VEE (TC-83) and Dengue 3 at 1ul of RNA stock per reaction. The results of these hybridizations are displayed in figures 17 and 18.

Figure 17

## Chip Design 3, Validation

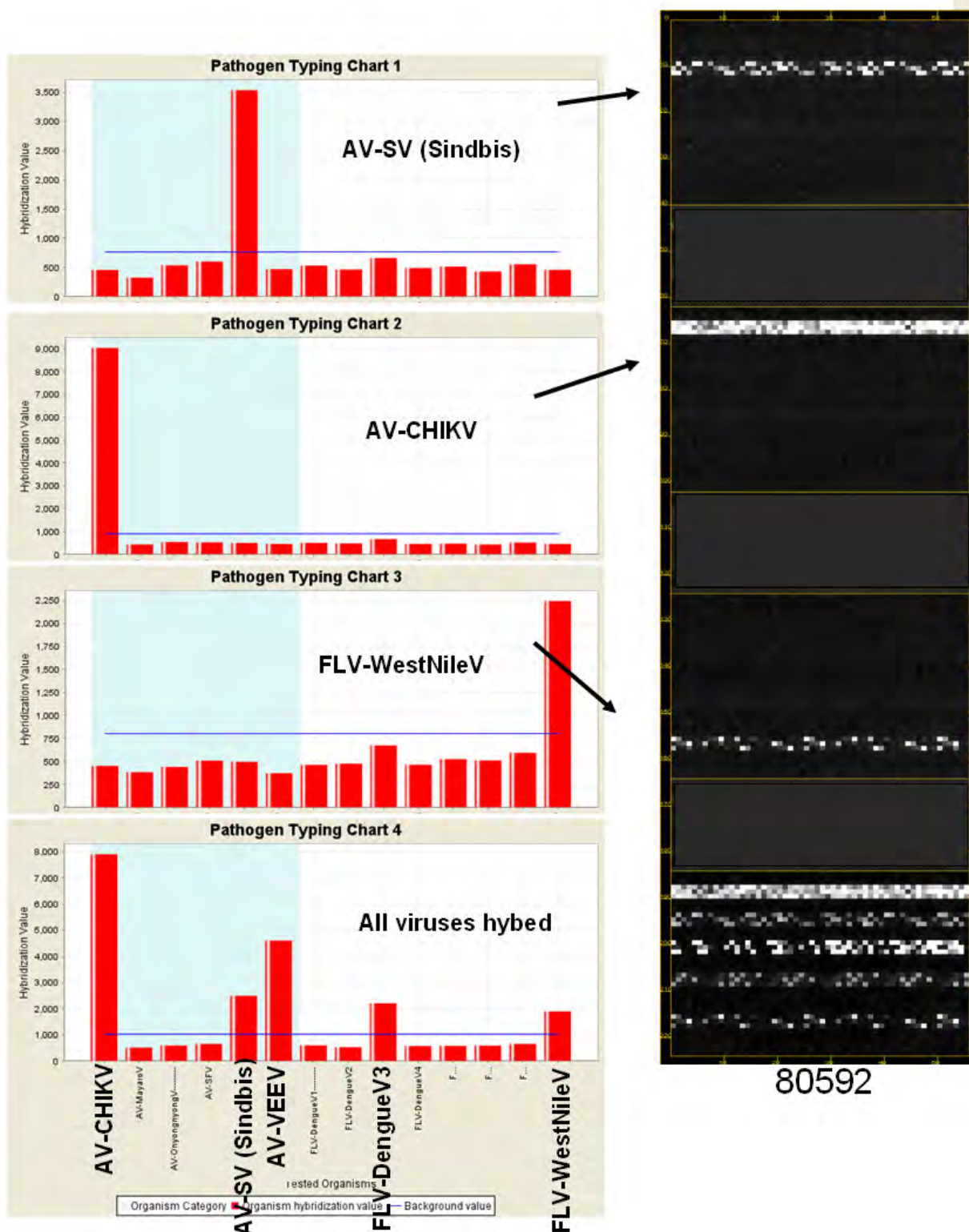
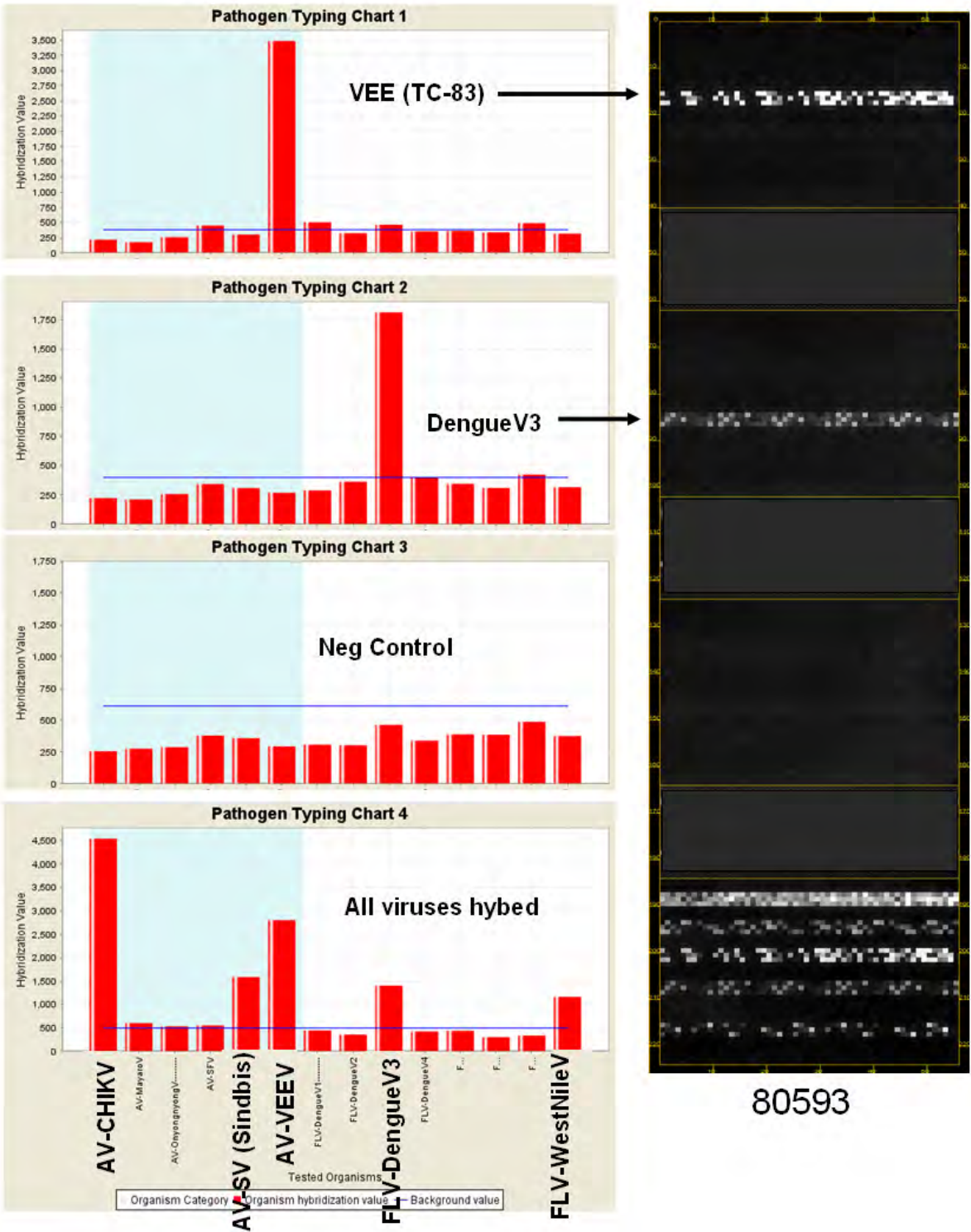


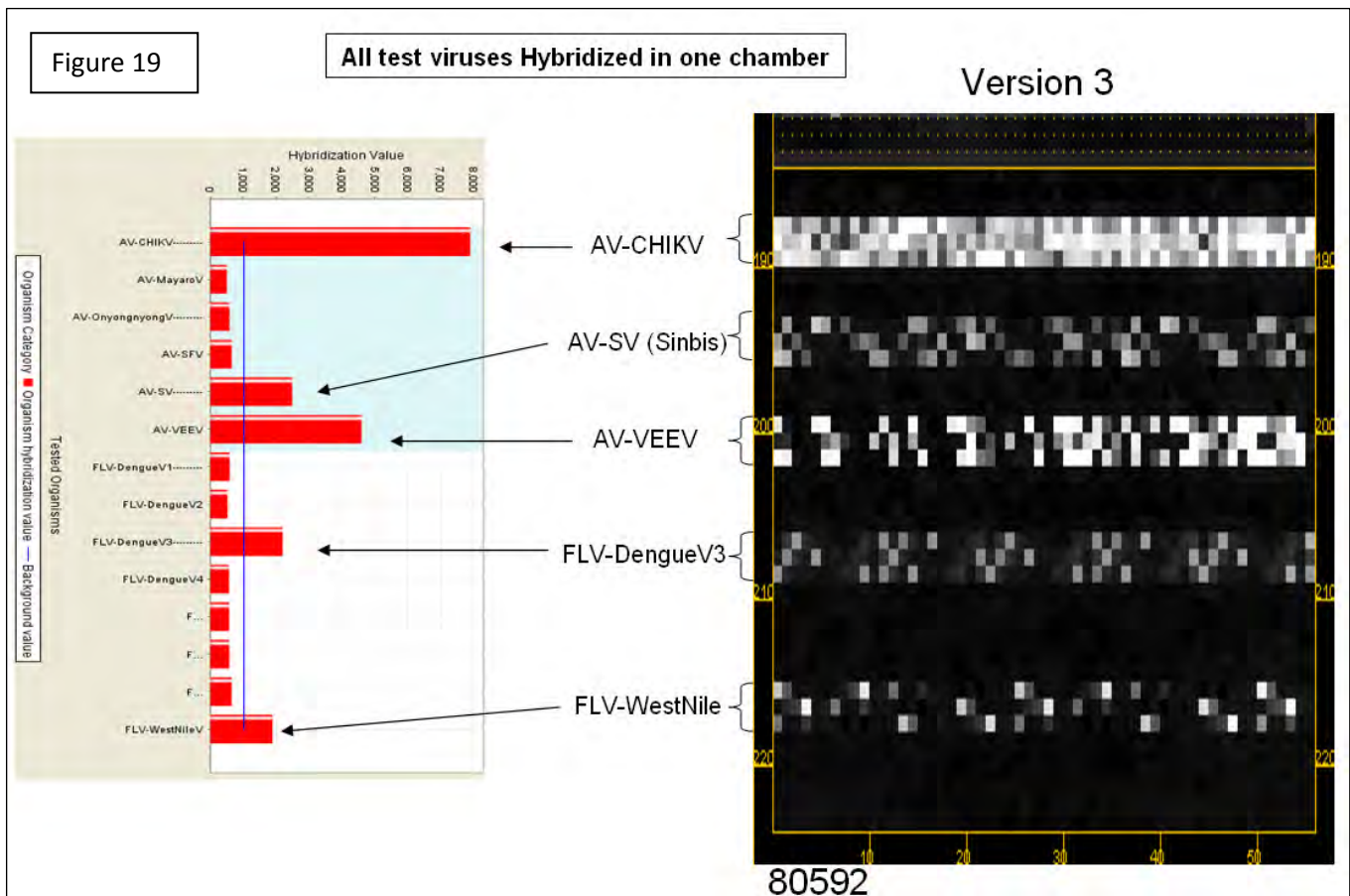
Figure 18

# Chip Design 3, Validation





The results of the new chip design, version #3, show an improved clarity of the detection from the version #1 chip design. Clearly, all viruses are detectable at high concentrations. It should be noted that chip 80593 showed signal of non-present viruses slightly above the blue “background” cut-off line. This is in spite of our natural ability to see the hybridization results and make a definitive call “by-eye” which tells us that only the VEE (tc-83) virus is present. Refinements to our pathogen calling algorithm, although minor, will be necessary for the final field deployable device. A close up sector 4, chip # 80952 (figure 17), displayed below as figure 19, shows all viruses hybed in one chamber.

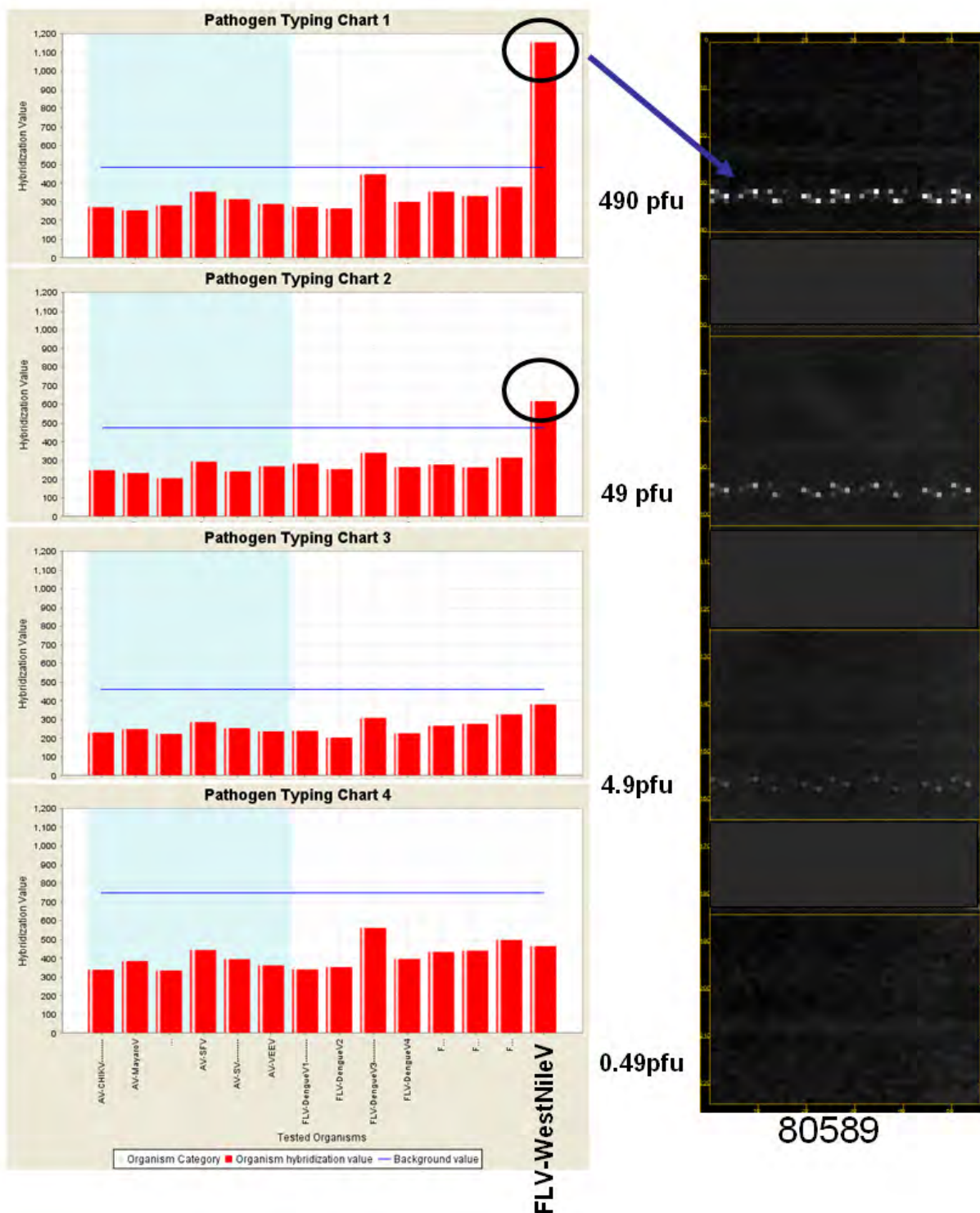


To further validate chip design #3, the Sindbis and WNV LOD titrations were hybridized to respective arrays. First, figure 20 shows the LOD hybridization for WNV where WNV is titrated 10-fold from 490 pfu per RT-PCR reaction in sector #1 to 0.49pfu (extinction dilution) in sector # 4. The ElectraSense data analysis calls WNV as detected as low as 49pfu. This is an improvement from chip design #1, which showed WNV present at these levels, but also called the Flavivirus Omsk Hemorrhagic Fever as being present due to cross-hybridization (NOT shown, please refer to report #5).

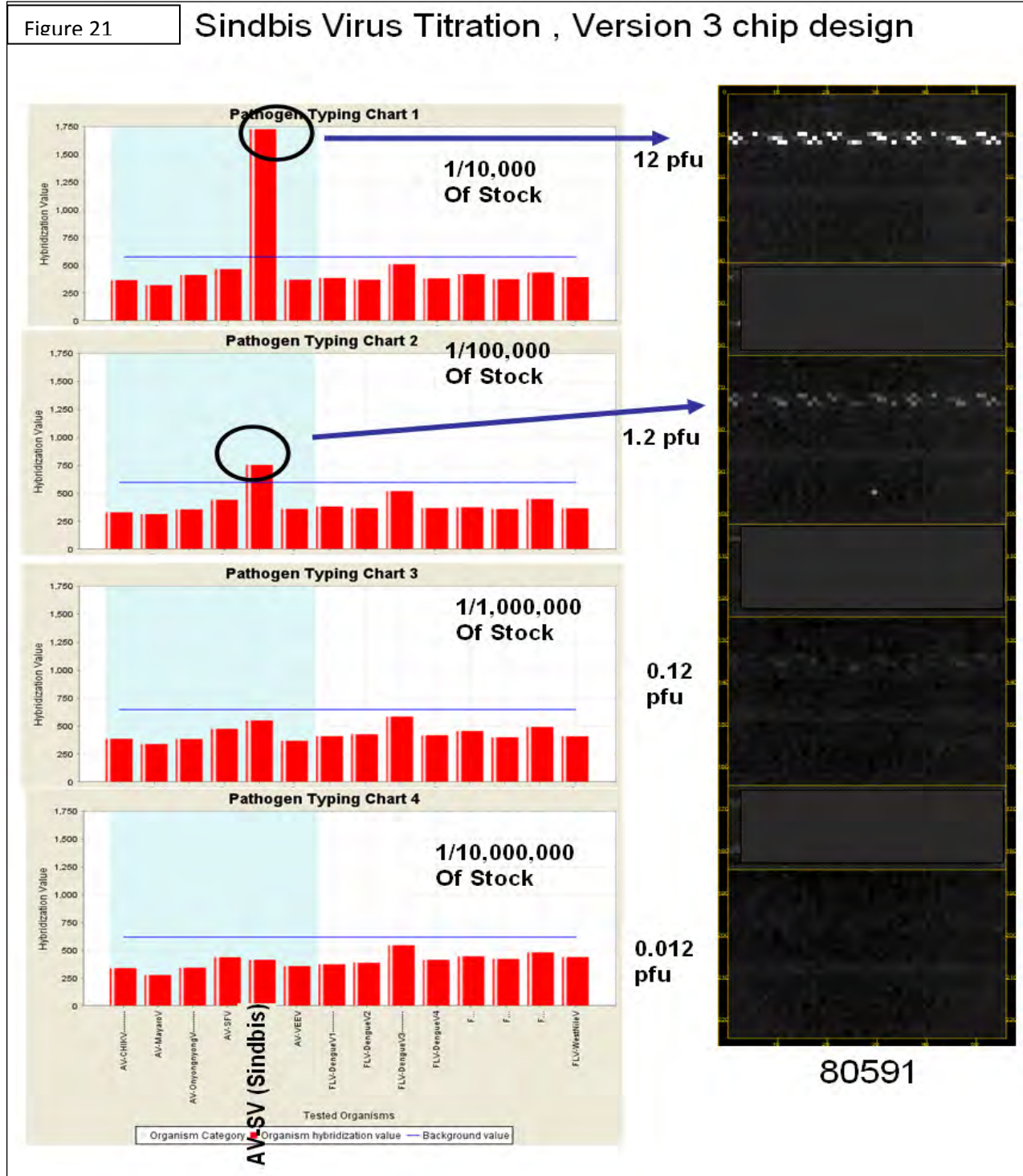


Figure 20

## West Nile virus Titration, Version 3 chip design



Secondly, figure 21 below shows the LOD hybridization for Sindbis, where Sindbis is titrated 10-fold from 12pfu per RT-PCR reaction in sector #1 to 0.012pfu (extinction dilution) in sector # 4. The ElectraSense data analysis calls Sindbis as detected as low as 1.2pfu. This is an improvement from chip design #1, which showed Sindbis present at only at 12 pfu, data not shown.



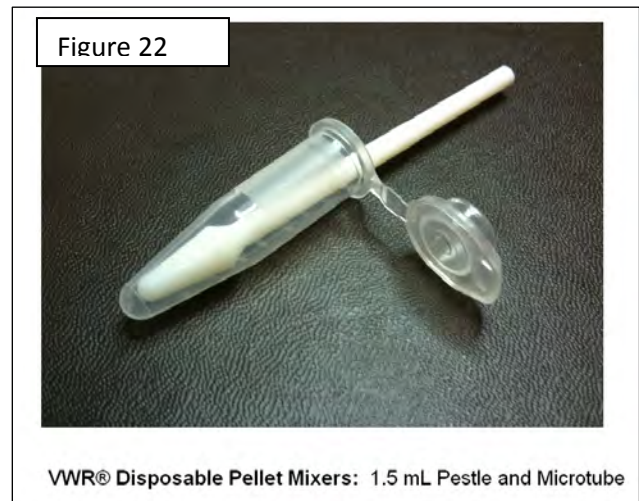
To summarize our iterative probe design and selection, we were able to refine the microarray chip design into a finalized revision, version#3, increasing our sensitivity by as much as 10-fold between version #1 and version #3. We used the version #3 as our final chip design for the prototype assay which will be tested as the deliverable for the phase I project, along with an ElectraSense II reader and all necessary detection reagents. This iterative process can be used for all viruses that will make up the ultimate assay, to detect all medically relevant arthropod borne viruses.

## Task 4: Sample Prep

Seattle Biomed supplied us with “cages” of mosquitoes (200 carcasses) for testing our grinding apparatus. These mosquitoes are uninfected Anopheles. Although not completely relevant, as most arborviruses are not vectored by anopheles, we used the carcasses in “mock” detection assays. The mock assay involves mixing the mosquito carcasses with a lysis solution, spiking-in viral RNA into this solutions, and purifying the spiked-in viral RNA from this solution mix. This mock assay will define our ability to extract RNA which will be usable in our amplification and detection strategy.

First, we tested our extraction protocol in which we have had success in the past grinding blue-green algae to extract RNA for gene expression studies. This grinding protocol is based on a pestle and mortar design, where the pestle is fabricated for insertion into a microfuge tube and grinding occurs by rotating the pestle inside the tube with the sample. The grinding takes place in the presence of a denaturant, in our case AVL buffer from the QIAamp Viral RNA Mini kit from Qiagen, and Diamond dust (1-10 mg). The diamond dust acts as an abrasive between the sample and the pestle and the wall of the microfuge tube. Unfortunately, this method was unsuccessful in producing RNA which was able to be RT-PCR amplified. We hypothesized that there may have been something created/released from the grind which carried over in the purification and inhibited the RT-PCR reaction.

To exclude problems with our in-house fabricated pestle, we decided to use a commercially available RNase/DNase free disposable pestle, shown in figure 22. We postulated that there could be problems from our in-house pestle/diamond-dust due to over-efficient grinding of the mosquitoes causing the release of a soluble inhibitor that co-purified with the RNA in the Qiagen column. We tested the VWR disposable pestle by grinding 20 to 30 mosquitoes in AVL buffer in the presence of spiked-in viral RNA. The grind was carried out by twisting the pestle, by hand, back and forth about 20 times, which created a black to brown solution, similar to our previous grinds with the in-house pestle, an example of which is shown at right, figure 23. Using this methodology we tested the following conditions:





Condition #2: 10ul of Sindbis stock spiked into the AVL/Mosquito mix, prior to grinding. Additional centrifugation step added to purification

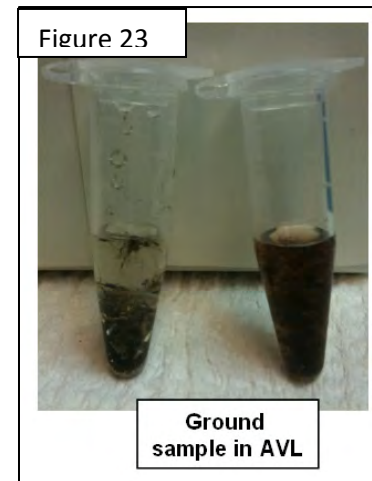
Condition #3: 10ul of Sindbis stock spiked into the AVL/Mosquito mix, prior to grinding.

Condition#4: 1ul of Sindbis stock spiked into the AVL/Mosquito mix, prior to grinding.

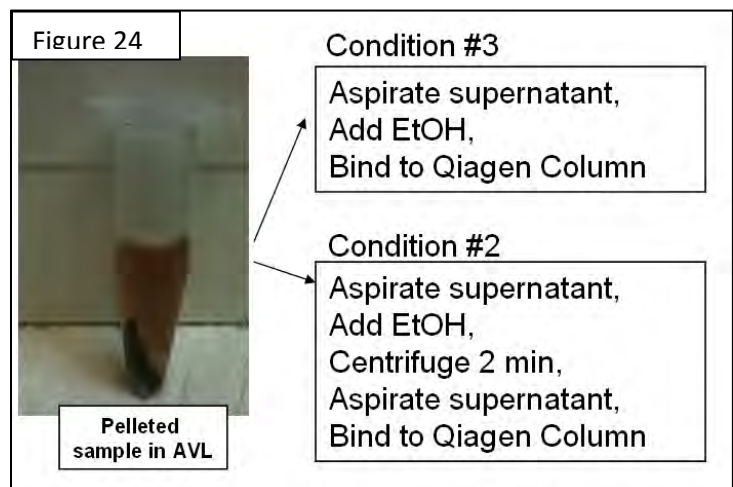
Condition#5: 10ul of Sindbis stock spiked into the AVL only mix, prior to grinding.

The protocol for grinding and the manipulation of the AVL lysate was performed as follows:

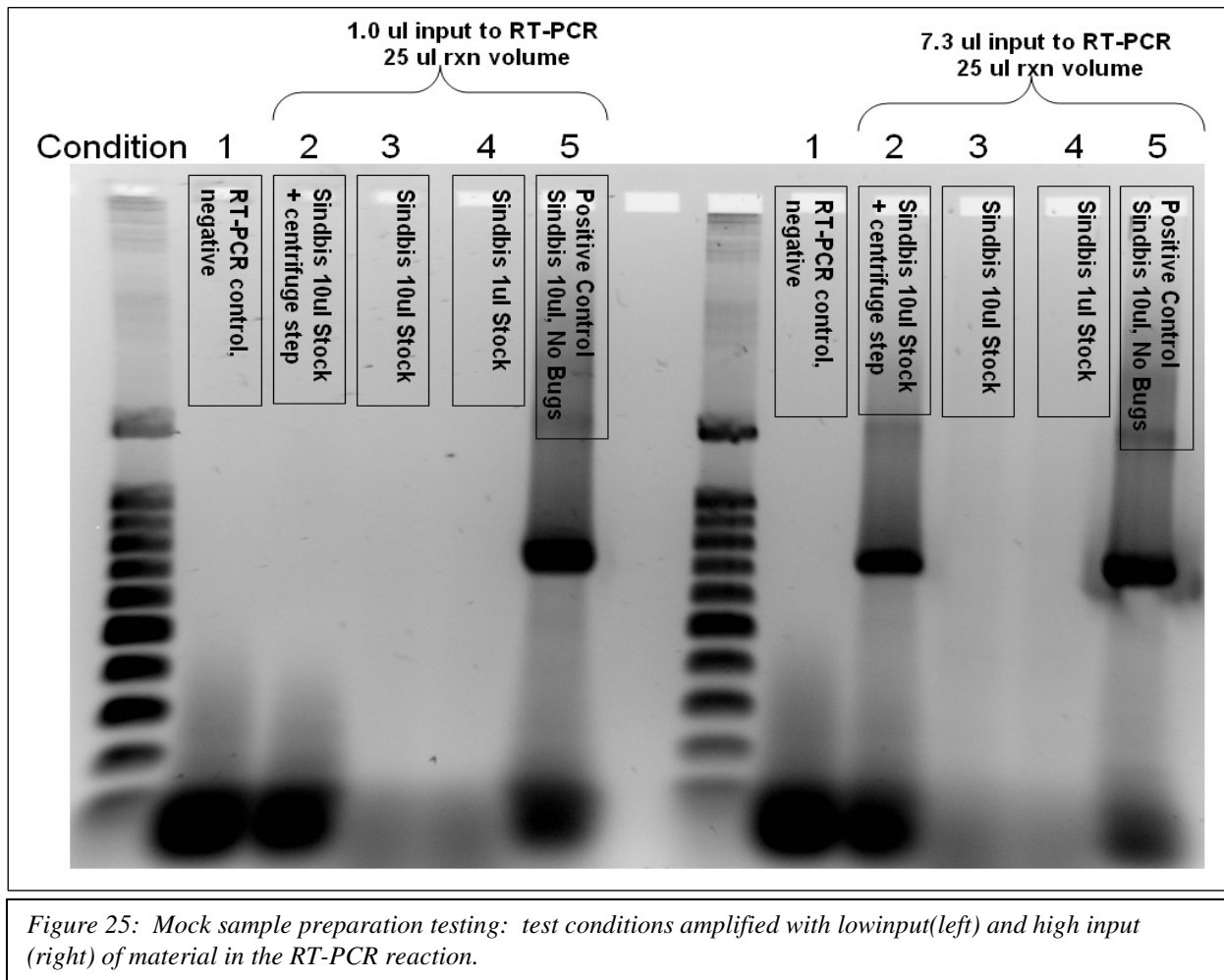
1. 280ul of AVL/Mosquito mix ground as described above using disposable pestle
2. 70ul of water added to mixture and centrifuged briefly to pellet particulate matter
3. \*pipet 250ul of this solution and mix with 200ul of EtOH. At this point, condition #2 varied from all other conditions in that there was an additional centrifugation step added to pellet any left over particulate from the AVL/EtOH mixture.
4. AVL/EtOH mixture pipetted onto Qiagen column and purification proceeded as per the manufacture protocol. RNA was eluted from the column with 60ul of AVE buffer, provided by Qiagen.



The major difference between condition #2 and #3 was the addition of a centrifugation step after mixing the AVL with EtOH, as shown in the picture at right, figure 24. It should be noted that the color of the AVL/EtOH solution looked the same between these two conditions, regardless of the centrifugation step. However, the Qiagen column from condtion#2 was not dark or discolored as in condition #3, and the final elution was clear from this column as well. All elutions from mosquito grinds had a slight charcoal/grey color, excluding condtion#2.



Elutions from these reactions were seeded into the RT-PCR amplification reaction. Two quantities were used as input to test two reactions per condition: 1ul and 7.3ul. 7.3ul was used as this is the maximum volume that is able to be used in the 25ul reaction. 7.5ul of the completed amplification was run out on a gel to visualize successful production of the Sindbis amplification fragment. Those results are shown in figure 25 below:



Firstly, the positive control verifies that we are able to take a small quantity of RNA, subject it to our grinding apparatus, and recover it with enough efficiency to detect it with RT-PCR. The Sindbis stock concentration is calculated to be  $1.2 \times 10^5$  pfu/ul. Using an input of 10ul, that would be  $1.2 \times 10^6$  viral genomes. Assuming 100% efficient recovery, the 60ul elution would contain all of these genomes, translating to a concentration of  $2 \times 10^4$  pfu/ul of Sindbis genomes. 1 and 7.3ul of this as input into an RT-PCR reaction translates to  $2 \times 10^4$  and  $1.46 \times 10^5$  copies of Sindbis genomes, respectively. Both of these conditions worked for the control purification which did not contain any exposure to mosquitoes.

Secondly, Condition #3 and #4 show no amplification, and the lack of primer tailing, or the small molecular weight product present in the negative and positive control conditions, argues for an inhibition of the RT-PCR reaction in these conditions. Condition #2, however showed amplification of the Sindbis fragment when 7.3ul of the eluted grind was used as input into the RT-PCR reaction. Condition #3 also shows amplification of the small product suggesting that there was no inhibition of the RT-PCR amplification in either the 1ul or 7.3ul reactions. We can conclude from these results that the extra centrifugation step used on condition #3 after the mixture of AVL with EtOH had an important effect on removing inhibitors of the RT-PCR reaction.

These results give us a reasonable launching board for further optimization of this grinding/purification procedure, and ultimately, will be the key to the development of a fully automated extraction system. Because we know that whatever is inhibiting the RT-PCR reaction can be centrifuged out of the system, we can translate the centrifugation step into a filter step and easily port this filter step onto an automated system, in the end.

### **Task 6: Deliverable**

We have assembled the deliverable for this phase I SBIR, which are 100 assays and a detection device. The 100 assays consist of 25 4x2k arrays of the version #3 array design and include enough multiplexed primer mix to perform the RT-PCR amplification which we have optimized, as well as an ElectraSense reader to perform the detection step. We also provided a full protocol and analysis software with these assays.

## Key Research Accomplishments from Phase I

### Tasks:

- 1. Bioinformatics:** The bioinformatics effort completed in this phase I period laid the ground work for future expansion of the assay to include all medically relevant arthropod-borne viruses. Using the database that we created along with the optimized design algorithms, we will be able to expand our prototype assay “in silico” to include all of the viruses listed in the FOA as well as others that may prove relevant, ultimately depending upon the field setting. This bioinformatics effort will, in the end, be the cornerstone for expanding the assay in a phase II effort.
- 2. Amplification strategy:** We demonstrated that our RT-PCR strategy for producing hybridizable material in a multiplexed fashion is both specific and sensitive. Using this strategy we demonstrated near single copy detection in LOD tests. Furthermore, we confirmed the ability of our RT-PCR strategy to co-amplify multiple viral species, simultaneously.
- 3. Alternate amplification strategy:** We piloted and demonstrated the feasibility of using a multiplexed padlock assay to replace RT-PCR as our amplification strategy.
- 4. Chip design testing and refinement, and assay validation:** Using the probes designed from our bioinformatics effort, we validated our chip design with hybridization of amplified material to our microarray platform. We confirmed detection of individual and mixed species in hybridization experiments, and we refined our design to increase the sensitivity and specificity of the overall assay. Ultimately we tested 3 chip design iterations to accomplish this task.
- 5. Sample Prep:** Through empirical testing, we discovered a method to purify spiked-in viral RNA from “mock” mosquito grinding and extraction assays. These observations give us a reasonable starting point for further optimization of this grinding/purification procedure, and ultimately, will be a key to the development of a fully automated extraction system.
- 6. Data Analysis:** A prototype data analysis pipeline was written to analyze multiple hybes in batch mode. This software can analyze the hybridization patterns and make a call based on probe signal intensities.

## **Reportable Outcomes**

Funding applied for based on work supported by this contract: We have submitted a Phase II proposal with The U.S. Army Medical Research and Materiel Command (USAMRMC) for the continued development and automation of our prototype assay.



## Conclusions

The overall goal of this project was to develop a prototype assay for the detection of viruses from class A arboviruses (alphaviruses) and class B arboviruses (flaviviruses). To this end, we developed a prototype of a detection system which includes: Probe design, sample prep, multiplexed PCR amplification, hybridization, detection, and data analysis. To demonstrate our prototype, we successfully completed all of the tasks that we outlined in our original phase I proposal. Ultimately, this research effort demonstrated the detection of 5 viral species, 2 Flaviviruses and 3 alphaviruses, with near single copy detection in a single reaction/detection assay. The detection device itself is a hand-held product that requires a USB connection, and uses that connection as its sole power source. Our phase I effort successfully lays the ground work for further developing this prototype assay into a fully automated field-deployable detection system.

To summarize the Phase I assay work-flow, we created an assay which included the following three steps: (1) Extract RNA from Mosquitoes; (2) Amplify and label viral RNA loci; and (3) Detect amplification product on microarray. The results of the assay were analyzed, using software we created, to make a pathogen call. With additional funding, it will be possible to further develop this prototype assay, and by expanding upon these three steps we will be able to (1) automate the sample prep system to extract and purify the RNA, requiring only an input of mosquitoes; (2) create an amplification system which will enable multiplexed amplification of all medically relevant arboviruses; and (3) automate the hybridization and electrochemically detection for hands free processing. We ultimately envision integrating the automated sample prep and the amplification system into one disposable cartridge which will insert into a device for processing. The output of this cartridge will be hybridization ready material appropriate for biosensor interrogation using a developed hybridizer/reader.

“So what section”

So what is the strength of the CustomArray platform? The strength of the platform is its core technology which is a semiconductor (CMOS) electrode array that is used to electrochemically synthesize oligonucleotides in-situ. This array of electrodes can also be used as multiplexed biosensors to detect binding events, which is the core of our electrochemical detection technology. Often times our array is pigeonholed as simply a DNA Microarray device, and the fact that it is an electrode array is forgotten. But it is much more. Viewed as an established multiplexed biosensor, the CustomArray platform is the ONLY platform of its kind that can enable field detection of multiple pathogens. Our system simply needs the necessary automation to bring the material to be interrogated to these sensors, which is the chip. Our phase II proposal, which we submitted prior to this report, outlines this automation effort.

There is a pressing need to detect pathogens, invasive species, and biothreat agents using a compact and robust assay device. By bringing robust and multiplexed detection into the field, we will be able to dramatically improve the discovery of emerging pathogenic threats. The CustomArray system is the only system that has the breadth and the redundancy to accurately detect multiple species in real world situations where novel forms are continually emerging.

## References

None

## Appendices

List of personnel receiving pay from the research effort:

John Cooper: Principal Investigator

Dominic Suciu: Bioinformatics lead

Marcelo Caraballo: Sr. Scientist

Sho Fuji: V.P. Engineering and Production

Todd West: Engineering Technician